

Risk aversion, health behaviour, and adverse selection in the German market for private supplementary health insurance

Hendrik Schmitz*

Ruhr Graduate School in Economics (RGS Econ)
and University of Duisburg-Essen

March 31, 2009

Abstract

I analyse the degree of private (i.e., unknown to the insurer) information in the German market for private supplementary insurance for hospital stays. Using data from the SOEP 2002-2007 I find a significant positive correlation between holding supplementary insurance and having a hospital stay, conditional on the information that is observed by the insurer. The degree of adverse selection, however, is only very small. The reason for this is that there are different sources of adverse and advantageous selection that partly net out. Non-smoking, income, and mental health are all sources of advantageous selection, because they increase the probability to buy insurance but decrease the likelihood of a hospital stay. The directly measured degree of risk aversion, however, is a source of adverse selection.

JEL Classification: D82, G22, I11

Keywords: adverse selection, advantageous selection, supplementary health insurance, risk aversion, health behaviour

*I am grateful to Matthias Keese, Harald Kühnemund, Annika Meng, Reinhold Schnabel and participants of the Fifth International Young Scholars SOEP Symposium for valuable comments. I thank Steffen John, and Caroline Schulz and for explaining how private insurance companies calculate their insurance premia and Sandra Schaffner for the list of risky jobs. All errors are my own. Financial support by the Leibniz association is gratefully acknowledged. All correspondence to Hendrik Schmitz, Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen), Hohenzollernstr. 1-3, 45128 Essen, Germany, Fax: +49-201-8149-500, e-mail: schmitz@rwi-essen.de

1 Introduction

Standard insurance models with asymmetric information like the Rothschild-Stiglitz-model predict a positive correlation of insurance cover and the occurrence of the insured risk conditional on the information of the insurance provider. That is, individuals who are bad risks choose insurance from a set of offered contracts that has a higher coverage than do good risks. Private individual information on the true risk type prevents insurance companies from perfectly calculating proper insurance premia for all risk types which might drive the good risks out of the market (i.e., there is adverse selection in the insurance market). An empirical test on the positive correlation between insurance cover and the occurrence of the risk (the “positive correlation test”) conditional on all the information that is observed by the insurance company can be seen as a test on the presence of asymmetric information in an insurance market (Chiappori and Salanie, 2001).

However, many empirical applications find no evidence for adverse selection in different insurance markets like markets for long-term care, Medigap insurance or life-insurance (see Cutler, Finkelstein, and McGarry (2008) for an overview). In some markets there can even be found a negative correlation between insurance cover and experience of risks. One explanation for this finding is that individuals do not only have private information on their risk type (possibly leading to adverse selection) but that also preferences like risk aversion shape their demand for insurance and the probability of the experience of a risk. A more risk-averse individual might demand more insurance but at the same time try to minimise the probability to experience the risk. Risk aversion is unobserved by the insurer but this information asymmetry does not lead to adverse selection, it is rather a source of the opposite, namely advantageous selection. Because the private information about the risk type and risk preferences is not one-dimensional but multi-dimensional, potential sources of adverse and advantageous selection may net out and the overall effect is not clear a priori.

Therefore, the positive correlation test as a test on asymmetric information is invalid (Finkelstein and Poterba, 2006). Failure to reject the null of no asymmetric information can arise if there are sources of adverse and advantageous selection that net out. Finkelstein and Poterba (2006) thus propose another test on information asymmetry. The existence of only one variable that is not used by the insurer to calculate the risk classification of the insured, which is correlated with both the insurance choice and the risk of the insured loss (conditional on the observed variables) is evidence for asymmetric information.

In this study I analyse the existence and direction of information asymmetries in the German market for private supplementary insurance for hospital stays. In Germany, about 90 per cent of the population are covered by public insurance that generally covers all the expenditures for hospital visits (except for a small amount of co-payments). However, additional to this basic insurance these individuals can buy private supplementary insurance that enables a better treatment during the hospital stay, namely a double room and treatment by a chief physician. When buying this private supplementary insurance, individuals have to give a detailed statement about their age, gender, kind of job and current health status (including a complete list of conditions). Hence, all these variables are known by the insurance companies and can thus be used to calculate the monthly insurance premium. However, several important variables are unobserved by the insurer. Not only might the true health status still be unknown (thus being a source of adverse selection), but so is also the degree of risk aversion (possibly being a source of advantageous selection). Risk aversion is usually not observed by the econometrician either, but the individual health behaviour is often used as a proxy.

The results indicate that there is a positive correlation of holding insurance and having a hospital stay, conditional on observed variables, i.e., there is adverse selection. However, the degree of adverse selection is rather small. This is because there are several unobserved variables like income, smoking, and the mental health state that work in the opposite direction (i.e., are a source of advantageous selection). The

directly measured overall degree of risk aversion, however, reinforces the adverse selection effect.

This paper is organised as follows. The next section gives a brief overview of the German health insurance system with its private supplementary insurance and shows which results can be expected from theory. Section 3 explains the data, sections 4 the estimation and testing method. Section 5 reports the results, while section 6 checks their robustness. Section 7 concludes.

2 Institutional background and expectations from theory

One standard insurance model incorporating asymmetric information was developed by Rothschild and Stiglitz (1976). In their model individuals have private information about their risk type, i.e., their propensity to suffer from the loss they seek to insure against. Individuals can only hold one insurance contract at a time and there is perfect competition among insurance providers. Rothschild and Stiglitz (1976) show that, depending on the share of good risks, either there is no equilibrium at all or an equilibrium in which the good risks buy less insurance than the bad risks. Private individual information on the true risk type prevents insurance companies from perfectly calculating proper insurance premia for all risk types which might drive the good risks out of the market (i.e., there is adverse selection in the insurance market). While Rothschild and Stiglitz (1976) assume that individuals only differ in their risk type, de Meza and Webb (2001) also allow for differences in risk preferences among the individuals. In their model, individuals do not only have private information about their risk type but also on their degree of risk aversion. Risk aversion, however, both affects the insurance coverage and the risk type. Risk averse individuals (named the “timid” by de Meza and Webb (2001)) demand more insurance and lower their risks by preventive behaviour, whereas the “bold” care less for insurance and prevention, thus increasing their risk. While the Rothschild-Stiglitz-model predicts a positive correlation of risk type and insurance coverage (the higher the risk, the more insurance), de Meza and

Webb (2001) allow for equilibria that exhibit a negative correlation between risk and insurance coverage.

Cutler, Finkelstein, and McGarry (2008) review the literature and find that whether there is more insurance of high-risk individuals or less insurance differs between insurance markets. While in acute care and annuity markets one usually finds that high-risk individuals buy more insurance (as the standard model also predicts) this is not the case in life insurance, long-term care, and Medigap-markets.

The insurance market I analyse in this paper is the German market for private supplementary insurance. In Germany, about 90 per cent of all individuals are covered by public health insurance (called the “statutory health insurance”). In general, this is a full cover insurance. Except for small co-payments for doctor visits, hospital visits, and prescription drugs it covers all health care expenditures caused by the insured. As regards hospital visits, the statutory health insurance is basic in a sense that it does not cover stays in a double room and treatments by the chief physician. However, statutorily insured can buy private supplementary insurance that covers the extra costs for this improved quality. While the general public insurance premium does not depend on the risk type (the social health insurance is funded by payroll taxes) the premium for the private supplementary insurance premium is risk-adjusted. When buying the supplementary insurance, individuals have to give a detailed statement about their age, gender, kind of job and health status (including a complete list of conditions and previous treatments). The basic insurance premium depends on the age and sex of the insured individual. Health problems increase the premium. Finally, insurance companies can reject applicants with too risky jobs, or further increase their insurance premium.

There are two main advantages in analysing the market for private supplementary insurance for hospital visits in Germany. First, while in general the German health insurance system is strictly regulated, this is one market in the health care sector that exhibits a true competition. The vast majority of publicly insured is statutorily insured and has no chance to opt out of the system and to decide about their degree of insur-

ance cover. The only free choice they have is whether or not to buy private supplementary insurance. The second reason caters to the difference between moral hazard and adverse selection. Moral hazard also results from the information asymmetry between insured and insurance provider and leads to a positive correlation between insurance and consumption of health care services, because with more insurance cover the price for health services decreases and, thus, the demand increases *ceteris paribus*. If the consumption of services and the experience of risk is measured in the same unit (e.g. health care expenditures, number of doctor visits, number of hospitalisations), it is difficult to disentangle moral hazard from adverse selection when a positive correlation is found. However, as regards hospital visits the price elasticity is usually found to be very low. I therefore assume that moral hazard does not play a role in the case of hospital visits and that therefore a positive correlation can solely be attributed to adverse selection.

3 Data

I use data from the German Socio-Economic Panel (SOEP), which is a large-scale panel data set.¹ It started in 1984 in West Germany and was extended to include East Germany in June 1990. There were several refreshments resulting in a sample size of more than 20,000 adult individuals living in more than 12,000 households that participated in the SOEP survey in 2006 (Wagner, Frick, and Schupp, 2007).

The SOEP includes information about the health insurance status (private or public). In case of public insurance it also asks whether or not an individual holds private supplementary insurance of several different types. One of these includes supplementary insurance for hospital visits. Concerning the second outcome variable, the occurrence of the insured risk, the SOEP contains a variable indicating at least one hospital

¹The data used in this paper were extracted using the Add-On package PanelWhiz v2.0 (Nov 2007) for Stata. PanelWhiz was written by Dr. John P. Haisken-DeNew (john@panelwhiz.eu).

The PanelWhiz generated DO file to retrieve the SOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are my own. Haisken-DeNew and Hahn (2006) describe PanelWhiz in detail.

visit in the previous twelve months. However, as a measure of risk I am not interested in hospital stays in the past but in the future probability of entering a hospital. Therefore I use the individual information from the next wave to form the variable *hospital* which takes on the value one in case of a hospital visit within one year *after* the interview and zero in case of no hospital visit.

The SOEP includes all the variables known by the insurer to calculate the insurance premium, except for the detailed list of conditions of the individuals. However, it contains the Physical Component Summary Scale (PCS), a measure of physical health which is formed by the SF12-questionnaire (see Nübling, Andersen, Mühlbacher, Schupp, and Wagner (2007) for a description) and the BMI as an objective measure of health. In addition, the degree of handicap is given in the data. These variables - together with the self-assessed health status - give a very informative picture of the individuals' health. In order to proxy the accident risk at the workplace (which affects the insurance premium), I generate a dummy variable *risky job* which takes on the value one if an individual works in one of the 20 jobs with the highest death rate. This holds for about 1 per cent of all jobs.²

Overall, the knowledge of the insurance companies is represented well in the data set. As the PCS and the BMI are only available in the waves of 2002, 2004, and 2006, I only use these three waves. After dropping all individuals with missing information in one of the covariates and those about 10 per cent holding private full cover insurance I can use 49,479 observations in person-year form resulting from 21,925 different individuals. 8.6 per cent hold private supplementary insurance for hospital visits, and 11.7 per cent had at least one hospital visit in the following year (see the descriptive statistics in table A1 in the appendix).

The SOEP includes several variables that are not observed by insurance companies and that are likely to affect both the insurance choice and the likelihood of entering

²This acts as a proxy for overall riskiness at the workplace. The list of jobs is taken from Schaffner and Spengler (2005) who use HVBG/BUK data on accidents in combination with the IAB-Beschäftigtenstichprobe. The data cover the period between 1985 and 1995. As only jobs are included in the list that are subject to social insurance contributions I also include risky jobs that are paid according to civil service law like police men, firefighters and soldiers.

a hospital. For instance, variables that indicate the individual health behaviour and which are often used as proxies for risk aversion. These are how frequent a person does sports, whether she smokes or not, or follows a health-conscious diet. All of these variables can be expected to have an effect on the health status, at least in the long-run. However, current values of these variables are not very informative as behavioural measures because they are highly endogenous. A recent hospital visit due to bad health often increases the health consciousness and lets individuals stop smoking, eat more healthy and do more sports. In fact, the self-assessed health status is negatively correlated with smoking and not being on a healthy diet.³ This should not be interpreted as a causal negative effect of healthy behaviour on health but as a reversed causality (from health status to healthy behaviour). To solve this problem I construct a variable that indicates if a person has ever smoked within the last ten years and the frequency of sports five years ago.⁴ Since information about a healthy diet is only available in the waves of 2004 and 2006, I only use the year 2006 and the variable healthy diet two years ago when focussing on this health behaviour as a source of advantageous selection. Using the lagged variables, I shall interpret them as general health behaviours that are not affected by the current health status but, the other way around, affect the current health status.

The SOEP also contains a direct measure of risk aversion. The question in the SOEP is: "How would you rate your willingness to take risks?" It is measured on an 11-point scale from 0 (not willing to take risks) to 10 (completely ready to take risks). This question was asked twice, in the years 2004 and 2006. I assume that the degree of risk aversion is a time invariant preference (at least over a period of some years) which, however, is measured with considerable error. To reduce the measurement error, I take the individual average value of risk aversion over the two years and assign each

³This is not so, however, for the correlation between self-assessed health and frequency of sports. All correlations not reported here.

⁴For the smoking variable I use information in the 2002 wave about the smoking history of the individuals. As regards the sports variable I do not take the value ten years ago because this would result in too many missing values (for all individuals that have not been in the sample ten years ago). Using the ten year lagged variable, however, does not change the results at all.

individual this constant risk aversion over the entire period (i.e. from 2002 to 2006).⁵

Finally, the SOEP contains a measure of mental health (again from the SF12-questionnaire) and the household income as variables which are observed by the individual but not by the insurer and which might effect both the demand for insurance and the hospital visits.

4 Methods

As a first step I perform one version of the positive correlation test as proposed by, e.g., Chiappori and Salanie (2001). Although Finkelstein and Poterba (2006) convincingly argue that the test is invalid, I use it as a test on the overall direction of information asymmetry and, thus, as a first indication. The test employs a bivariate probit with the choice to buy supplementary insurance and the occurrence of a hospital stay in the following year as the two left hand side variables. On the right hand side appear only the variables which are observed by the insurance companies (health status, age, gender, and job risk; all included in the vector X in equations (1) to (4)). To account for non-linearities in the effects of the health status, I also include several interactions between the variables.

$$HOSPITAL = 1[X\beta_1 + \mu_1 > 0] \quad (1)$$

$$INSURANCE = 1[X\delta_1 + \mu_2 > 0] \quad (2)$$

The sign and significance of ρ (which is the estimated correlation between μ_1 and μ_2) then informs about asymmetric information and whether there is adverse or advantageous selection. Note that the estimated coefficients in this bivariate probit do

⁵Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005) show in an experimental setting with a pre-test group of the SOEP households that the risk aversion measure is reliable, although being self-assessed. In 2004 there is also a question about the risk aversion concerning one's own health. As with the health behaviour this variable is endogenous with the ill individuals indicating a higher degree of risk aversion concerning their own health. Therefore I abstract from this variable.

not measure causal effects of the included variables on the probability of buying insurance and entering a hospital because they are likely to suffer from omitted variable bias. The aim of this test, however, is not to find causal effects but to estimate the correlation in the unobserved part that remains when all the information available by the insurance company is controlled for.

However, ρ just measures the overall degree of asymmetric information and there might still be sources of adverse and advantageous selection that cancel out in the market for private supplementary insurance. Thus, in the second step, these kinds of sources are analysed further. Therefore, I add one of the unobserved variables (i.e., unobserved by the insurer; named B in equations (3) and (4)) at a time to the two probit equations to analyse their impact on the insurance choice and the risk occurrence.

$$HOSPITAL = 1[X\beta_1 + \beta_2 B + \mu_1 > 0] \quad (3)$$

$$INSURANCE = 1[X\delta_1 + \delta_2 B + \mu_2 > 0] \quad (4)$$

Having a positive (or negative) sign of both β_2 and δ_2 then reveals that the variable is associated with a higher (lower) demand for insurance and a higher (lower) risk of occurrence. Because this variable is not observed by the insurance company, this is a source of adverse selection. A variable that has opposing signs of β_2 and δ_2 , however, is a source of advantageous selection as it increases the demand for insurance and decreases the likelihood of experiencing a risk at the same time (or vice versa).

5 Results

The results of the bivariate probit are reported in table 1. Because the health variables are strongly interacted and because of the non-linearity of the model, the single coefficients are somewhat difficult to interpret. Furthermore they do not have a causal interpretation but just measure associations. However, the bivariate probit is mainly

carried out in order to perform the positive correlation test. Therefore, I restrict the interpretation to the estimated correlation of the error terms.

Table 1: Bivariate probit estimation

	Private suppl. insurance		Hospital Visit	
Female	0.222	(0.30)	0.234	(0.19)
Age	0.040***	(0.01)	-0.026***	(0.01)
Age squared/1000	-0.270***	(0.06)	0.240***	(0.04)
PCS	-0.782	(2.01)	-4.273***	(1.32)
PCS squared	1.652	(1.55)	0.553	(1.15)
PCS x Age	-0.003	(0.01)	0.019**	(0.01)
PCS x Female	-0.434	(0.42)	0.546**	(0.26)
PCS x BMI	-0.004	(0.02)	0.028**	(0.01)
Female x Age	0.001	(0.00)	-0.008***	(0.00)
BMI x Age	-0.000	(0.00)	-0.000	(0.00)
BMI high	-0.048	(0.05)	0.020	(0.04)
BMI very high	0.017	(0.09)	-0.058	(0.07)
SAH very good	0.124*	(0.07)	-0.096	(0.07)
SAH good	0.032	(0.04)	-0.038	(0.04)
SAH poor or bad	-0.136**	(0.06)	0.288***	(0.04)
Disabled	0.046	(0.12)	-0.010	(0.07)
Degree Disability	-0.001	(0.00)	0.003***	(0.00)
Risky Job	0.322**	(0.13)	0.098	(0.12)
Constant	-2.195***	(0.72)	0.562	(0.45)
ρ	0.075***	(0.03)		
Observations	49,479			

*p<0.1, **p<0.05, ***p<0.01; Standard errors in parentheses, clustered by individuals; all regressions weighted

The results indicate that there is a small but significantly positive correlation between insurance coverage and hospital visits conditional on the information known by the insurer (estimated $\rho = 0.075$), thus indicating adverse selection. That is, there are some unobserved factors that both increase the probability of a hospital visit and the likelihood of holding supplementary insurance. As noted in the beginning, this is only the aggregate degree of information asymmetry. Table 2 shows the regression results when one possible source of adverse or advantageous selection is added to the probit equations at a time. Here, only the estimated parameters of the added variables are shown and discussed, i.e., each row represents one regression. The full regression

Table 2: Sources of asymmetric information

	Insurance	Hospital	Observations
Smoked in prev. 10 years	-0.082*	0.059**	45,037
No healthy food 2 years ago	-0.097*	-0.080	13,248
Sports 5 years ago	0.108***	0.000	26,849
Household income	0.134***	-0.017**	47,842
Mental health	0.468***	-0.563***	49,478
Risk aversion	0.051***	0.016**	48,243

*p<0.1, **p<0.05, ***p<0.01; All regressions including covariates observed by insurers;
all regressions weighted

results can be found in table A2 and A3 in the appendix.

The results indicate that there are several sources of advantageous selection in the German market for private supplementary insurance. Individuals who have smoked in the previous ten years tend to buy less supplementary insurance. At the same time, however, they have a higher probability of a hospital stay in the next year. Smokers seem to care less for their own health (or they seem to be more risk loving). Because smoking is not observed by the insurer it is a source of advantageous selection. Never following a health-conscious diet also decreases the probability of buying supplementary insurance but does not decrease the probability of hospital stays. The coefficient in the hospital equation is even negative (albeit not significant) indicating that eating healthy food increases the probability of a hospital visit. As mentioned earlier, this variable might suffer from a reversed causality with the less healthy caring more for healthy food. Apparently the two year lag of the variable is not long enough to solve the reversed causality problem.

The effect of doing sports on the probability of a hospital visit is zero. But individuals who exercise frequently also tend to buy more insurance.

Like smoking, income and the mental health status are also sources of advantageous selection. Individuals with more income and a better state of mental health buy more supplementary insurance and have less hospital visits at the same time. The first result is in line with Fang, Keane, and Silverman (2008) who find that individuals with higher income are more likely to hold Medigap insurance and to spend less money on

medical care.

According to the direct measure of risk aversion, risk loving individuals tend to buy more private supplementary insurance. Note that this is a measure of overall risk aversion and not just financial risk aversion (as usually - if at all - asked in general surveys) or risk aversion concerning health. It can be argued that a lower general risk aversion increases the probability of occurrence of a loss and that individuals use their private information about their risk type when deciding to buy or not to buy supplementary insurance. Indeed, risk loving individuals also have more hospital stays. That is, contrary to the results of the proxy variables for risk aversion (the variables indicating the health behaviour), according to the direct measure, risk aversion seems to be a source of adverse selection.

6 Robustness checks

The positive correlation test indicated that there is a low degree of adverse selection in the German market for private supplementary insurance. Although the employed health measures give a complete picture of the individual health, one might argue that the SOEP does include worse information on the health status than the insurance companies actually have. Therefore, the test might falsely find evidence of adverse selection which is in fact only due to missing information on the true health status in the SOEP. To check the robustness of the result I draw on another data set that includes a detailed list of subjective and objective health measures, namely the 1st wave of the Survey of Health, Aging and Retirement in Europe (SHARE), a survey of the population aged 50+ conducted in 2004. The survey involved 19,286 households and 32,022 individuals, covering a wide range of topics, including physical health, behaviour, socioeconomic status, and income. I use the subset of German respondents which add up to 2,438 individuals holding public insurance (after dropping individuals with missing values in one variable). As health measures I include self-assessed health as well as the number of chronic conditions, the number of reported symptoms, the number

of problems with ADLs, the number of problems with IADLs, and the grip strength. The main drawback of the SHARE data set for this analysis is the small sample size of the German sub sample. Since only individuals above the age of 50 are included and, moreover, the dependent variable is the incidence of a hospital visit in the previous year and not in the following year, the results of the tests are not fully comparable.⁶ The number of individuals holding private supplementary insurance is similar to the one in the SOEP file (9.0 per cent) whereas the incidence of a hospital visit in the last year is higher (15.2 per cent) due to the higher average age of the respondents. Table 3 reports the results. Again, the test finds a remaining positive correlation of more insurance and hospital visits, which is a sign of adverse selection. The size of ρ is even slightly higher than in the SOEP data set, however, due to a much lower number of observations not as precisely estimated. All in all, however, the inclusion of even more detailed information on the individual health status does not change the results and makes the effects found in the previous section reliable.

Several other specifications were estimated - now again with the original SOEP data set - to check the robustness of the results. First of all, the definition of the experience of the occurred risk was changed. Because a hospital visit is in general a rare case, I extended the binary variable to "hospital visits within the next two years" resulting in a one for about 19 per cent of all sample members compared to 11 per cent before. Unfortunately, as the most recent wave is the one of 2007 and I need two "future" waves to construct the variable, I cannot use the information from the 2006 wave for the analysis, leading to a loss of about one third of the sample size. The results, however, are robust to this change.⁷

Finally, I estimated the same regressions for males and females separately (see table 4). Basically the results from the last section hold for both groups, although some estimated parameters are not significant anymore (in part due to a smaller sample

⁶When this study was performed, the second wave of the SHARE, representing the year 2006, was already released. However, for the variable "hospital visit in two years", which is more comparable to the outcome variable "hospital visit in the next year" that I use in the SOEP, only individuals that have information in both years could be used. This resulted in a panel size of only about 1,300 individuals making significant results even harder to find than in the above used sample.

⁷Not shown here but available upon request.

Table 3: Bivariate probit estimation - SHARE data set

	Private suppl. insurance		Hospital Visit	
Female	0.391	(0.56)	0.501	(0.48)
Age	0.053	(0.05)	0.038	(0.04)
Age squared/1000	-0.539	(0.36)	-0.234	(0.30)
Number of chronic conditions	-0.014	(0.04)	-0.001	(0.00)
Number of symptoms	-0.007	(0.03)	0.092***	(0.02)
BMI	-0.138*	(0.07)	0.000	(0.00)
Female x Age	-0.005	(0.01)	-0.010	(0.01)
BMI x Age	0.001	(0.00)	-0.000	(0.00)
Grip Strength	0.008*	(0.00)	-0.007	(0.00)
Number ADL	0.100	(0.10)	0.023	(0.06)
Number IADL	-0.143	(0.09)	-0.025	(0.06)
SAH very good	0.483***	(0.14)	-0.565***	(0.15)
SAH good	0.233**	(0.10)	-0.376***	(0.08)
SAH poor or bad	-0.074	(0.16)	0.309***	(0.10)
Constant	-1.674	(2.45)	-2.190	(1.38)
Rho	0.096	(0.06)		
Observations		2438		

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Standard errors in parentheses, clustered by individuals; all regressions weighted

size). The most remarkable difference is that smoking has no effects in the females sample and the overall effect is solely driven by the males.

7 Conclusion

Using data from the SOEP I find evidence for adverse selection in the German market for private supplementary insurance for hospital stays. However, the degree of adverse selection is very small. The reason is that there are different sources of adverse and advantageous selection that partly cancel out. A detailed analysis using data that are unobserved by the insurance company and, thus, private to the insured, reveals that smoking, income, and mental health are all sources of advantageous selection. People who often do sports and eat more healthy also have a higher likelihood of holding supplementary insurance. However, there seem to be no (short-term) positive health effects of this behaviour. The directly measured degree of risk aversion,

Table 4: Sources of asymmetric information - Males and Females

		Insurance	Hospital	Observations
Males	Risk aversion	0.051***	0.022**	22,176
	Smoked in prev. 10 years	-0.137**	0.064	20,629
	No healthy food 2 years ago	0.069	-0.028	6,007
	Sports 5 years ago	0.104***	0.005	12,250
	Household income	0.134***	-0.020	22,027
	Mental health	0.153	-0.549***	22,768
Females	Risk aversion	0.045***	0.010*	26,438
	Smoked in prev. 10 years	-0.036	0.029	24,760
	No healthy food 2 years ago	-0.126***	-0.017	7,350
	Sports 5 years ago	0.133***	0.003	14,906
	Household income	0.124***	-0.005	26,184
	Mental health	0.681***	-0.667***	27,091

*p<0.1, **p<0.05, ***p<0.01; All regressions including covariates observed by insurers; all regressions weighted

however, seems to be a source of adverse selection.

How does this finding match the prediction of de Meza and Webb (2001)? They show that risk aversion can be a source of advantageous selection. Our results imply that this is true for one of our proxy variables of risk aversion, namely smoking. Smokers can be seen as more risk loving (concerning their health) who buy less insurance and have a higher likelihood of needing a hospital visit. The direct measure of risk aversion, however, works in the opposite direction. One explanation for this is that risk aversion is also not a one- but a multi-dimensional variable. Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2005) show that people have different degrees of risk aversion in different areas like risk aversion concerning financial matters, health, or driving. These preferences are all correlated but may have different shapes. Here, we did not use the risk aversion concerning health matters due to econometric problems but the overall degree of risk aversion which is somewhat different. People seem to have their degree of risk aversion in mind when they are planning to buy supplementary insurance and use this private information.

Finally, what does the finding of several sources of adverse and advantageous selection imply? Were the insurance companies able to observe all the variables, they

could calculate proper insurance premia for all individuals with their different characteristics. However, as they are not, the information asymmetry prevents a first best outcome of the market equilibrium. Even if there is one-dimensional private information, the market outcome is inefficient. The inefficiency even increases if private information is multi-dimensional as is the case here (see Finkelstein and McGarry (2006) for a discussion).

Acknowledgement

This paper uses data from SHARE Waves 1 & 2, as of December 2008. SHARE data collection in 2004-2007 was primarily funded by the European Commission through its 5th and 6th framework programmes (project numbers QLK6-CT-2001- 00360; RII-CT- 2006-062193; CIT5-CT-2005-028857). Additional funding by the US National Institute on Aging (grant numbers U01 AG09740-13S2; P01 AG005842; P01 AG08291; P30 AG12815; Y1-AG-4553-01; OGHA 04-064; R21 AG025169) as well as by various national sources is gratefully acknowledged (see <http://www.share-project.org> for a full list of funding institutions).

References

- CHIAPPORI, P.-A., AND B. SALANIE (2001): "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 108(1), 56–78.
- CUTLER, D. M., A. FINKELSTEIN, AND K. MCGARRY (2008): "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance," *American Economic Review*, 98(2), 157–62.
- DE MEZA, D., AND D. C. WEBB (2001): "Advantageous Selection in Insurance Markets," *RAND Journal of Economics*, 32(2), 249–62.
- DOHMEN, T., A. FALK, D. HUFFMAN, U. SUNDE, J. SCHUPP, AND G. G. WAGNER (2005): "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," Iza discussion papers, Institute for the Study of Labor (IZA).

- FANG, H., M. P. KEANE, AND D. SILVERMAN (2008): "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, 116(2), 303–350.
- FINKELSTEIN, A., AND K. MCGARRY (2006): "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *American Economic Review*, 96(4), 938–958.
- FINKELSTEIN, A., AND J. POTERBA (2006): "Testing for Adverse Selection with "Unused Observables"," Nber working papers, National Bureau of Economic Research, Inc.
- HAIKEN-DENEW, J. P., AND M. HAHN (2006): "PanelWhiz: a flexible modularized Stata interface for accessing large-scale panel data sets," .
- NÜBLING, M., H. H. ANDERSEN, A. MÜHLBACHER, J. SCHUPP, AND G. G. WAGNER (2007): "Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of SF12v2," *Journal of Applied Social Science (Schmollers Jahrbuch)*, 127(1), 171–182.
- ROTHSCHILD, M., AND J. E. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90(4), 630–49.
- SCHAFFNER, S., AND H. SPENGLER (2005): "Der Einfluss unbeobachteter Heterogenität auf kompensatorische Lohndifferentiale und den Wert eines Statistischen Lebens: Eine mikroökonomische Parallelanalyse mit IABS und SOEP," Darmstadt Discussion Papers in Economics 152, Institut für Volkswirtschaftslehre (Department of Economics), Technische Universität Darmstadt (Darmstadt University of Technology).
- WAGNER, G. G., J. R. FRICK, AND J. SCHUPP (2007): "The German Socio-Economic Panel Study (SOEP): scope, evolution, and enhancements," *Journal of Applied Social Science (Schmollers Jahrbuch)*, 127(1), 139–169.

Appendix

Table A1: Descriptive statistics of sample variables

Variable	Mean	Std. Dev.	Min	Max	Observations
Supplementary insurance	0.086	0.281	0	1	49478
Hospital visit next year	0.117	0.322	0	1	49478
Risk aversion	4.493	2.057	0	10	48242
Smoked in prev. 10 years	0.383	0.486	0	1	45036
No healthy food 2 years ago	0.505	0.500	0	1	13248
Sports 5 years ago	2.019	1.236	1	4	26849
Household income	2.743	1.887	0	101.799	47841
Mental health	0.499	0.101	0.019	0.778	49478
Age	47.475	17.425	16	99	49478
Physical health	0.495	0.101	0.092	0.781	49478
25 < BMI < 30	0.357	0.479	0	1	49478
BMI > 30	0.155	0.362	0	1	49478
Self rated health very good	0.092	0.290	0	1	49478
Self rated health good	0.403	0.491	0	1	49478
Self rated health bad or poor	0.171	0.377	0	1	49478
Handicapped	0.122	0.327	0	1	49478
Degree of handicap	7.055	20.808	0	100	49478
Risky job	0.010	0.098	0	1	49478

Table A2: Probit results - Insurance

	(1)	(2)	(3)	(4)	(5)	(6)
Risk aversion	0.051*** (0.01)					
Smoked in prev. 10 years		-0.082* (0.04)				
No healthy food 2 years ago			-0.097* (0.06)			
Sports 5 years ago				0.108*** (0.02)		
Household income					0.134*** (0.01)	
Mental health						0.468*** (0.18)
Age	0.042*** (0.01)	0.041*** (0.01)	0.050*** (0.02)	0.015 (0.05)	0.028** (0.01)	0.041*** (0.01)
Age squared/1000	-0.261*** (0.06)	-0.281*** (0.07)	-0.342*** (0.10)	-0.242 (0.21)	-0.178*** (0.06)	-0.272*** (0.06)
PCS	-0.704 (2.05)	-0.790 (2.13)	1.817 (3.48)	-2.563 (5.14)	-1.458 (2.07)	-0.698 (1.99)
PCS squared	1.636 (1.56)	1.907 (1.60)	0.006 (3.00)	4.781 (3.42)	1.983 (1.59)	1.926 (1.53)
PCS x Age	-0.005 (0.02)	-0.006 (0.02)	-0.024 (0.02)	0.010 (0.04)	0.000 (0.02)	-0.004 (0.01)
PCS x Female	-0.427 (0.43)	-0.307 (0.45)	-0.407 (0.64)	-0.678 (0.92)	-0.208 (0.45)	-0.416 (0.42)
PCS x BMI	0.000 (0.02)	-0.008 (0.02)	-0.004 (0.03)	-0.051 (0.05)	0.002 (0.02)	-0.004 (0.02)
Female x Age	0.001 (0.00)	0.001 (0.00)	-0.001 (0.00)	0.007 (0.01)	0.002 (0.00)	0.001 (0.00)
BMI x Age	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)
BMI high	-0.050 (0.05)	-0.045 (0.05)	-0.115 (0.08)	-0.047 (0.11)	-0.047 (0.05)	-0.053 (0.05)
BMI very high	0.020 (0.09)	0.042 (0.09)	-0.174 (0.14)	0.187 (0.18)	0.041 (0.09)	0.012 (0.09)
SAH very good	0.097 (0.07)	0.123 (0.08)	0.179 (0.14)	0.097 (0.19)	0.033 (0.08)	0.057 (0.07)
SAH good	0.015 (0.04)	0.018 (0.04)	0.052 (0.07)	-0.017 (0.08)	-0.011 (0.04)	-0.006 (0.04)
SAH poor or bad	-0.109* (0.06)	-0.127** (0.06)	-0.016 (0.10)	-0.199* (0.11)	-0.095 (0.06)	-0.088 (0.06)
Disabled	0.009 (0.13)	0.093 (0.13)	0.051 (0.18)	0.182 (0.23)	0.030 (0.13)	0.041 (0.12)
Degree Disability	0.000 (0.00)	-0.001 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)
Risky Job	0.298** (0.13)	0.332** (0.14)	0.287 (0.20)	0.246 (0.30)	0.332** (0.14)	0.317** (0.13)
Female	0.238 (0.31)	0.140 (0.32)	0.242 (0.44)	-0.018 (0.74)	0.059 (0.32)	0.218 (0.30)
Constant	-2.577*** (0.76)	-2.159*** (0.79)	-3.037*** (1.11)	-1.558 (2.25)	-2.129*** (0.76)	-2.526*** (0.74)
Observations	48,243	45,037	13,248	26,849	47,842	49,478

*p<0.1, **p<0.05, ***p<0.01; Standard errors in parentheses, clustered by individuals;
all regressions weighted

Table A3: Probit results - Hospital visit

	(1)	(2)	(3)	(4)	(5)	(6)
Risk aversion	0.016*** (0.01)					
Smoked in prev. 10 years		0.059** (0.03)				
No healthy food 2 years ago			-0.080 (0.05)			
Sports 5 years ago				0.000 (0.02)		
Household income					-0.017** (0.01)	
Mental health						-0.563*** (0.12)
Age	-0.024*** (0.01)	-0.027*** (0.01)	-0.057*** (0.02)	-0.077*** (0.03)	-0.024*** (0.01)	-0.027*** (0.01)
Age squared/1000	0.224*** (0.04)	0.254*** (0.05)	0.436*** (0.09)	0.493*** (0.13)	0.224*** (0.04)	0.245*** (0.04)
PCS	-4.378*** (1.34)	-3.444** (1.38)	-3.680 (2.68)	-9.515*** (2.72)	-4.065*** (1.34)	-4.682*** (1.32)
PCS squared	0.724 (1.17)	0.162 (1.20)	-0.892 (2.29)	3.231 (2.39)	0.459 (1.17)	0.562 (1.14)
PCS x Age	0.018* (0.01)	0.014 (0.01)	0.033* (0.02)	0.058** (0.02)	0.017* (0.01)	0.021** (0.01)
PCS x Female	0.438* (0.27)	0.609** (0.28)	0.818 (0.52)	0.534 (0.51)	0.641** (0.27)	0.520** (0.26)
PCS x BMI	0.028** (0.01)	0.019 (0.01)	0.000 (0.02)	0.027 (0.03)	0.026* (0.01)	0.027** (0.01)
Female x Age	-0.008*** (0.00)	-0.007*** (0.00)	-0.007** (0.00)	-0.007* (0.00)	-0.007*** (0.00)	-0.008*** (0.00)
BMI x Age	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)	-0.000 (0.00)
BMI high	0.019 (0.04)	0.029 (0.04)	0.057 (0.07)	-0.014 (0.07)	0.032 (0.04)	0.025 (0.04)
BMI very high	-0.048 (0.07)	-0.033 (0.07)	0.087 (0.13)	-0.177 (0.12)	-0.038 (0.07)	-0.054 (0.07)
SAH very good	-0.099 (0.07)	-0.119* (0.07)	-0.061 (0.15)	-0.141 (0.19)	-0.109 (0.07)	-0.018 (0.07)
SAH good	-0.039 (0.04)	-0.043 (0.04)	-0.042 (0.08)	-0.010 (0.07)	-0.037 (0.04)	0.003 (0.04)
SAH poor or bad	0.287*** (0.04)	0.264*** (0.04)	0.308*** (0.08)	0.252*** (0.07)	0.284*** (0.04)	0.226*** (0.04)
Disabled	-0.025 (0.07)	-0.002 (0.08)	-0.102 (0.14)	-0.167 (0.12)	-0.012 (0.07)	-0.005 (0.07)
Degree Disability	0.003*** (0.00)	0.003** (0.00)	0.005** (0.00)	0.005*** (0.00)	0.003*** (0.00)	0.003** (0.00)
Risky Job	0.105 (0.12)	0.080 (0.12)	0.234 (0.25)	0.317 (0.25)	0.082 (0.12)	0.104 (0.12)
Female	0.334* (0.20)	0.133 (0.20)	0.047 (0.38)	0.189 (0.42)	0.166 (0.19)	0.243 (0.19)
Constant	0.435 (0.47)	0.204 (0.49)	1.526* (0.90)	3.317*** (1.19)	0.542 (0.46)	1.038** (0.47)
Observations	48,243	45,037	13,248	26,849	47,842	49,478

*p<0.1, **p<0.05, ***p<0.01; Standard errors in parentheses, clustered by individuals;
all regressions weighted